

## SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: Somerville, Chris  
Broun, Pierre  
van de Loo, Frank  
Boddupalli, Sekhar S.
- (ii) TITLE OF INVENTION: Production of Hydroxylated Fatty Acids in Genetically Modified Plants
- (iii) NUMBER OF SEQUENCES: 15
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: PILLSBURY MADISON & SUTRO  
(B) STREET: 1100 NEW YORK AVENUE, N.W.  
(C) CITY: WASHINGTON  
(D) STATE: D.C.  
(E) COUNTRY: USA  
(F) ZIP: 20005-3918
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: 3.5 inch, 1.44 MB storage  
(B) COMPUTER: IBM compatible  
(C) OPERATING SYSTEM: DOS 5.0  
(D) SOFTWARE: Word Perfect 5.1
- (vi) CURRENT APPLICATION DATA;  
(A) APPLICATION NUMBER: not yet assigned  
(B) FILING DATE: February 6, 1997  
(C) CLASSIFICATION:

## (2) INFORMATION FOR SEQ ID NO:1

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 543 nucleotides  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATTGGCACCCGGCGGCACCA TTCCAAACAAT GGATCCCTAG 40  
AAAAAGATGA AGTCTTGTC CCACCTAAGA AAGCTGCAGT 80  
CANATGGTAT GTCAAATAACC TCAACAAACCC TCTTGGACGC 120  
ATTCTGGTGT TAACAGTTCA GTTTATCCTC GGGTGGCCTT 160

Sequence Listing

*Sekhar  
A1*

TGTATCTAGC	CTTTAATGTA	TCAGGTAGAC	CTTATGATGG	200
TTTCGCTTCA	CATTTCTTCC	CTCATGCACC	TATCTTTAAG	240
GACCGTGAAC	GTCTCCAGAT	ATACATCTCA	GATGCTGGTA	280
TTCTAGCTGT	CTGTTATGGT	CTTTACCGTT	ACGCTGCTTC	320
ACAAGGATTG	ACTGCTATGA	TCTGCGTCTA	CGGAGTACCG	360
CTTTTGATAG	TGAACCTTTT	CCTTGTCTTG	GTCACTTTCT	400
TGCAGCACAC	TCATCCTTCA	TTACCTCACT	ATGATTCAAC	440
CGAGTGGGAA	TGGATTAGAG	GAGCTTTGGT	TACGGTAGAC	480
AGAGACTATG	GAATCTTGAA	CAAGGTGTTT	CACAACATAA	520
CAGACACCCA	CGTAGCACAC	CAC		543

## (2) INFORMATION FOR SEQ ID NO:2

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 544 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TATAGGCACC	GGAGGCACCA	TTCCAACACA	GGATCCCTCG	40
AAAGAGATGA	AGTATTGTC	CCAAAGCAGA	AATCCGCAAT	80
CAAGTGGTAC	GGCGAACATACC	TCAACAACCC	TCCTGGTCCG	120
ATCATGATGT	TAACTGTCCA	GTTCGTCTC	GGATGGCCCT	160
TGTACTTAGC	CTTCAACGTT	TCTGGCAGAC	CCTACAATGG	200
TTTCGCTTCC	CATTTCTTCC	CCAATGCTCC	TATCTACAAC	240
GACCGTGAAC	GCCTCCAGAT	TTACATCTCT	GATGCTGGTA	280
TTCTAGCCGT	CTGTTATGGT	CTTTACCGTT	ACGCTGTTGC	320
ACAAGGACTA	GCCTCAATGA	TCTGTCTAAA	CGGAGTTCCG	360
CTTCTGATAG	TTAACTTTT	CCTCGTCTTG	ATCACTTACT	400

*Sub  
A  
CMY*

TACAACACAC	TCACCCCTGCG	TTGCCTCACT	ATGATTCA	440
AGAGTGGGAT	TGGCTTAGAG	GAGCTTAGC	TACTGTAGAC	480
AGAGACTATG	GAATCTTGAA	CAAGGTGTT	CATAACATCA	520
CAGACACCCA	CGTCGCACAC	CACT		544

## (2) INFORMATION FOR SEQ ID NO:3

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1855 nucleotides
  - (B) TYPE: nucleotide
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAAGCTTT	ATAAGAAGTT	AGTTTTCTCT	GGTGACAGAG	40
AAATTNTGTC	AATTGGTAGT	GACAGTTGAA	GCAACAGGAA	80
CAACAAAGGAT	GGTTGGTGNT	GATGCTGATG	TGGTGATGTG	120
TTATTTCATCA	AATACTAAAT	ACTACATTAC	TTGTTGCTGC	160
CTACTTCTCC	TATTCCTCC	GCCACCCATT	TTGGACCCAC	200
GANCCTTCCA	TTTAAACCCCT	CTCTCGTGCT	ATTCAACCAGA	240
AGAGAAAGCCA	AGAGAGAGAG	AGAGAGAATG	TTCTGAGGAT	280
CATTGTCTTC	TTCATCGTTA	TTAACGTAAG	TTTTTTTGAA	320
CCACTCATAT	CTAAAATCTA	GTACATGCAA	TAGATTAATG	360
ACTGTTCCCTT	CTTTGATAT	TTTCAGCTTC	TTGAATTCAA	400
GATGGGTGCT	GGTGGAAAGAA	TAATGGTTAC	CCCCTCTTCC	440
AAGAAATCAG	AAACTGAAGC	CCTAAAACGT	GGACCATGTG	480
AGAAACCACC	ATTCACTGTT	AAAGATCTGA	AGAAAGCAAT	520
CCCACAGCAT	TGTTTCAAGC	GCTCTATCCC	TCGTTCTTTC	560
TCCTACCTTC	TCACAGATAT	CACTTTAGTT	TCTTGCTTCT	600
ACTACGTTGC	CACAAATTAC	TTCTCTCTTC	TTCCCTCAGCC	640

*Sue  
A  
cont.*

TCTCTCTACT	TACCTAGCTT	GGCCTCTCTA	TTGGGTATGT	680
CAAGGCTGTG	TCTTAACCGG	TATCTGGGTC	ATTGCCATG	720
AATGTGGTCA	CCATGCATTC	AGTGAECTATC	AATGGGTAGA	760
TGACACTGTT	GGTTTTATCT	TCCATTCCCTT	CCTTCTCGTC	800
CCTTACTTCT	CCTGGAAATA	CAGTCATCGT	CGTCACCATT	840
CCAACAATGG	ATCTCTCGAG	AAAGATGAAG	TCTTGTCCC	880
ACCGAAGAAA	GCTGCAGTCA	AATGGTATGT	TAATACCTC	920
AACAACCCTC	TTGGACGCAT	TCTGGTGTAA	ACAGTTCACT	960
TTATCCTCGG	GTGGCCTTG	TATCTAGCCT	TTAATGTATC	1000
AGGTAGACCT	TATGATGGTT	TCGCTTCACA	TTTCTTCCCT	1040
CATGCACCTA	TCTTTAAAGA	CCGAGAACGC	CTCCAGATAT	1080
ACATCTCAGA	TGCTGGTATT	CTAGCTGTCT	GTTATGGTCT	1120
TTACCGTTAC	GCTGCTTCAC	AAGGATTGAC	TGCTATGATC	1160
TGCGTCTATG	GAGTACCGCT	TTTGATAGTG	AACTTTTCC	1200
TTGTCTTGGT	AACTTTCTTG	CAGCACACTC	ATCCTTCGTT	1240
ACCTCATTAT	GATTCAACCG	AGTGGGAATG	GATTAGAGGA	1280
GCTTTGGTTA	CGGTAGACAG	AGACTATGGA	ATATTGAACA	1320
AGGTGTTCCA	TAACATAACA	GACACACATG	TGGCTCATCA	1360
TCTCTTGCA	ACTATACCGC	ATTATAACGC	AATGGAAGCT	1400
ACAGAGGCGA	TAAAGCCAAT	ACTTGGTGAT	TACTACCACT	1440
TCGATGGAAC	ACCGTGGTAT	GTGGCCATGT	ATAGGGAAAGC	1480
AAAGGAGTGT	CTCTATGTAG	AACCGGATAC	GGAACGTGGG	1520
AAGAAAGGTG	TCTACTATTA	CAACAATAAG	TTATGAGGCT	1560
GATAGGGCGA	GAGAAGTGCA	ATTATCAATC	TTCATTTCCA	1600
TGTTTTAGGT	GTCTTGTAA	AGAAGCTATG	CTTTGTTCA	1640
ATAATCTCAG	AGTCCATNTA	GTTGTGTCT	GGTGCATTTT	1680

*Sub  
A  
Cont'*

GCCTAGTTAT	GTGGTGTGCGG	AAGTTAGTGT	TCAAACGTGCT	1720
TCCTGCTGTG	CTGCCAGTG	AAGAACAAAGT	TTACGTGTTT	1760
AAAATACTCG	GAACGAATTG	ACCACAANAT	ATCCAAAACC	1800
GGCTATCCGA	ATTCCATATC	CGAAAACCGG	ATATCCAAAT	1840
TTCCAGAGTA	CTTAG			1855

## (2) INFORMATION FOR SEQ ID NO:4

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Gly	Ala	Gly	Gly	Arg	Ile	Met	Val	Thr
							5		10
Pro	Ser	Ser	Lys	Lys	Ser	Glu	Thr	Glu	Ala
							15		20
Leu	Lys	Arg	Gly	Pro	Cys	Glu	Lys	Pro	Pro
							25		30
Phe	Thr	Val	Lys	Asp	Leu	Lys	Lys	Ala	Ile
							35		40
Pro	Gln	His	Cys	Phe	Lys	Arg	Ser	Ile	Pro
							45		50
Arg	Ser	Phe	Ser	Tyr	Leu	Leu	Thr	Asp	Ile
							55		60
Thr	Leu	Val	Ser	Cys	Phe	Tyr	Tyr	Val	Ala
							65		70
Thr	Asn	Tyr	Phe	Ser	Leu	Leu	Pro	Gln	Pro
							75		80
Leu	Ser	Thr	Tyr	Leu	Ala	Trp	Pro	Leu	Tyr
							85		90
Trp	Val	Cys	Gln	Gly	Cys	Val	Leu	Thr	Gly
							95		100

*Sub  
a  
cont.*

Ile Trp Val Ile Gly His Glu Cys Gly His  
105 110

His Ala Phe Ser Asp Tyr Gln Trp Val Asp  
115 120

Asp Thr Val Gly Phe Ile Phe His Ser Phe  
125 130

Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr  
135 140

Ser His Arg Arg His His Ser Asn Asn Gly  
145 150

Ser Leu Glu Lys Asp Glu Val Phe Val Pro  
155 160

Pro Lys Lys Ala Ala Val Lys Trp Tyr Val  
165 170

Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ile  
175 180

Leu Val Leu Thr Val Gln Phe Ile Leu Gly  
185 190

Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser  
195 200

Gly Arg Pro Tyr Asp Gly Phe Ala Ser His  
205 210

Phe Phe Pro His Ala Pro Ile Phe Lys Asp  
215 220

Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp  
225 230

Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu  
235 240

Tyr Arg Tyr Ala Ala Ser Gln Gly Leu Thr  
245 250

Ala Met Ile Cys Val Tyr Gly Val Pro Leu  
255 260

Leu Ile Val Asn Phe Phe Leu Val Leu Val  
265 270

CONFIDENTIAL

*Jack  
AI  
CMT*

Thr Phe Leu Gln His Thr His Pro Ser Leu  
 275 280  
 Pro His Tyr Asp Ser Thr Glu Trp Glu Trp  
 285 290  
 Ile Arg Gly Ala Leu Val Thr Val Asp Arg  
 295 300  
 Asp Tyr Gly Ile Leu Asn Lys Val Phe His  
 305 310  
 Asn Ile Thr Asp Thr His Val Ala His His  
 315 320  
 Leu Phe Ala Thr Ile Pro His Tyr Asn Ala  
 325 330  
 Met Glu Ala Thr Glu Ala Ile Lys Pro Ile  
 335 340  
 Leu Gly Asp Tyr Tyr His Phe Asp Gly Thr  
 345 350  
 Pro Trp Tyr Val Ala Met Tyr Arg Glu Ala  
 355 360  
 Lys Glu Cys Leu Tyr Val Glu Pro Asp Thr  
 365 370  
 Glu Arg Gly Lys Lys Gly Val Tyr Tyr Tyr  
 375 380  
 Asn Asn Lys Leu

## (2) INFORMATION FOR SEQ ID NO:5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gly Gly Gly Gly Arg Met Ser Thr Val  
 5 10  
 Ile Thr Ser Asn Asn Ser Glu Lys Lys Gly  
 15 20

*Sub A  
cont*

Gly Ser Ser His Leu Lys Arg Ala Pro His  
25 30

Thr Lys Pro Pro Phe Thr Leu Gly Asp Leu  
35 40

Lys Arg Ala Ile Pro Pro His Cys Phe Glu  
45 50

Arg Ser Phe Val Arg Ser Phe Ser Tyr Val  
55 60

Ala Tyr Asp Val Cys Leu Ser Phe Leu Phe  
65 70

Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr  
75 80

Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp  
85 90

Leu Val Tyr Trp Leu Phe Gln Gly Cys Ile  
95 100

Leu Thr Gly Leu Trp Val Ile Gly His Glu  
105 110

Cys Gly His His Ala Phe Ser Glu Tyr Gln  
115 120

Leu Ala Asp Asp Ile Val Gly Leu Ile Val  
125 130

His Ser Ala Leu Leu Val Pro Tyr Phe Ser  
135 140

Trp Lys Tyr Ser His Arg Arg His His Ser  
145 150

Asn Ile Gly Ser Leu Glu Arg Asp Glu Val  
155 160

Phe Val Pro Lys Ser Lys Ser Lys Ile Ser  
165 170

Trp Tyr Ser Lys Tyr Ser Asn Asn Pro Pro  
175 180

Gly Arg Val Leu Thr Leu Ala Ala Thr Leu  
185 190

*Sub A  
cont*

Leu Leu Gly Trp Pro Leu Tyr Leu Ala Phe  
195 200

Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe  
205 210

Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile  
215 220

Phe Ser Glu Arg Glu Arg Leu Gln Ile Tyr  
225 230

Ile Ala Asp Leu Gly Ile Phe Ala Thr Thr  
235 240

Phe Val Leu Tyr Gln Ala Thr Met Ala Lys  
245 250

Gly Leu Ala Trp Val Met Arg Ile Tyr Gly  
255 260

Val Pro Leu Leu Ile Val Asn Cys Phe Leu  
265 270

Val Met Ile Thr Tyr Leu Gln His Thr His  
275 280

Pro Ala Ile Pro Arg Tyr Gly Ser Ser Glu  
285 290

Trp Asp Trp Leu Arg Gly Ala Met Val Thr  
295 300

Val Asp Arg Asp Tyr Gly Val Leu Asn Lys  
305 310

Val Phe His Asn Ile Ala Asp Thr His Val  
315 320

Ala His His Leu Phe Ala Thr Val Pro His  
325 330

Tyr His Ala Met Glu Ala Thr Lys Ala Ile  
335 340

Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr  
345 350

Asp Gly Thr Pro Phe Tyr Lys Ala Leu Trp  
355 360

Sub  
At  
Cost

Arg Glu Ala Lys Glu Cys Leu Phe Val Glu  
365 370  
Pro Asp Glu Gly Ala Pro Thr Gln Gly Val  
375 380  
Phe Trp Tyr Arg Asn Lys Tyr  
385

## (2) INFORMATION FOR SEQ ID NO:6

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Ala Gly Gly Arg Met Pro Val Pro  
5 10  
Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr  
15 20  
Thr Lys Arg Val Pro Cys Glu Lys Pro Pro  
25 30  
Phe Ser Val Gly Asp Leu Lys Lys Ala Ile  
35 40  
Pro Pro His Cys Phe Lys Arg Ser Ile Pro  
45 50  
Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile  
55 60  
Ile Ile Ala Ser Cys Phe Tyr Tyr Val Ala  
65 70  
Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro  
75 80  
Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp  
85 90  
Ala Cys Gln Gly Cys Val Leu Thr Gly Ile  
95 100  
Trp Val Ile Ala His Glu Cys Gly His His  
105 110

Sub  
al  
un

Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp  
115 120

Thr Val Gly Leu Ile Phe His Ser Phe Leu  
125 130

Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser  
135 140

His Arg Arg His His Ser Asn Thr Gly Ser  
145 150

Leu Glu Arg Asp Glu Val Phe Val Pro Lys  
155 160

Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys  
165 170

Tyr Leu Asn Asn Pro Leu Gly Arg Ile Met  
175 180

Met Leu Thr Val Gln Phe Val Leu Gly Trp  
185 190

Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly  
195 200

Arg Pro Tyr Asp Gly Phe Ala Cys His Phe  
205 210

Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg  
215 220

Glu Arg Leu Gln Ile Tyr Leu Ser Asp Ala  
225 230

Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr  
235 240

Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser  
245 250

Met Ile Cys Leu Tyr Gly Val Pro Leu Leu  
255 260

Ile Val Asn Ala Phe Leu Val Leu Ile Thr  
265 270

Tyr Leu Gln His Thr His Pro Ser Leu Pro  
275 280

DRAFT Sequence

*Sub  
Al  
Cont*

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu  
 285 290  
 Arg Gly Ala Leu Ala Thr Val Asp Arg Asp  
 295 300  
 Tyr Gly Ile Leu Asn Lys Val Phe His Asn  
 305 310  
 Ile Thr Asp Thr His Val Ala His His Leu  
 315 320  
 Phe Ser Thr Met Pro His Tyr Asn Ala Met  
 325 330  
 Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu  
 335 340  
 Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro  
 345 350  
 Trp Tyr Val Ala Met Tyr Arg Glu Ala Lys  
 355 360  
 Glu Cys Ile Tyr Val Glu Pro Asp Arg Glu  
 365 370  
 Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn  
 375 380  
 Asn Lys Leu

## (2) INFORMATION FOR SEQ ID NO:7

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Gly Ala Gly Gly Arg Met Gln Val Ser  
 5 10  
 Pro Pro Ser Lys Lys Ser Glu Thr Asp Asn  
 15 20  
 Ile Lys Arg Val Pro Cys Glu Thr Pro Pro  
 25 30

Sub A Cont

Phe Thr Val Gly Glu Leu Lys Lys Ala Ile  
35 40

Pro Pro His Cys Phe Lys Arg Ser Ile Pro  
45 50

Arg Ser Phe Ser His Leu Ile Trp Asp Ile  
55 60

Ile Ile Ala Ser Cys Phe Tyr Tyr Val Ala  
65 70

Thr Thr Tyr Phe Pro Leu Leu Pro Asn Pro  
75 80

Leu Ser Tyr Phe Ala Trp Pro Leu Tyr Trp  
85 90

Ala Cys Gln Gly Cys Val Leu Thr Gly Val  
95 100

Trp Val Ile Ala His Glu Cys Gly His Ala  
105 110

Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp  
115 120

Thr Val Gly Leu Ile Phe His Ser Phe Leu  
125 130

Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser  
135 140

His Arg Arg His His Ser Asn Thr Gly Ser  
145 150

Leu Glu Arg Asp Glu Val Phe Val Pro Arg  
155 160

Arg Ser Gln Thr Ser Ser Gly Thr Ala Ser  
165 170

Thr Ser Thr Thr Phe Gly Arg Thr Val Met  
175 180

Leu Thr Val Gln Phe Thr Leu Gly Trp Pro  
185 190

Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg  
195 200

*Sub  
A  
Cont*

Pro Tyr Asp Gly Gly Phe Ala Cys His Phe  
205 210

His Pro Asn Ala Pro Ile Tyr Asn Asp Arg  
215 220

Glu Arg Leu Gln Ile Tyr Ile Ser Asp Ala  
225 230

Gly Ile Leu Ala Val Cys Tyr Gly Leu Leu  
235 240

Pro Tyr Ala Ala Val Gln Gly Val Ala Ser  
245 250

Met Val Cys Phe Leu Arg Val Pro Leu Leu  
255 260

Ile Val Asn Gly Phe Leu Val Leu Ile Thr  
265 270

Tyr Leu Gln His Thr His Pro Ser Leu Pro  
275 280

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu  
285 290

Arg Gly Ala Leu Ala Thr Val Asp Arg Asp  
295 300

Tyr Gly Ile Leu Asn Gln Gly Phe His Asn  
305 310

Ile Thr Asp Thr His Glu Ala His His Leu  
315 320

Phe Ser Thr Met Pro His Tyr His Ala Met  
325 330

Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu  
335 340

Gly Glu Tyr Tyr Gln Phe Asp Gly Thr Pro  
345 350

Val Val Lys Ala Met Trp Arg Glu Ala Lys  
355 360

Glu Cys Ile Tyr Val Glu Pro Asp Arg Gln  
365 370

*Sub  
at  
Cent*

Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn  
375 380

Asn Lys Leu Xaa

(2) INFORMATION FOR SEQ ID NO:8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 309 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ser Leu Leu Thr Ser Phe Ser Tyr Val Val  
5 10

Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr  
15 20

Ile Ala Thr Thr Tyr Phe His Leu Leu Pro  
25 30

Gln Pro Phe Ser Leu Ile Ala Trp Pro Ile  
35 40

Tyr Trp Val Leu Gln Gly Cys Leu Leu Thr  
45 50

Arg Val Cys Gly His His Ala Phe Ser Lys  
55 60

Tyr Gln Trp Val Asp Asp Val Val Gly Leu  
65 70

Thr Leu His Ser Thr Leu Leu Val Pro Tyr  
75 80

Phe Ser Trp Lys Ile Ser His Arg Arg His  
85 90

His Ser Asn Thr Gly Ser Leu Asp Arg Asp  
95 100

Glu Arg Val Lys Val Ala Trp Phe Ser Lys  
105 110

Tyr Leu Asn Asn Pro Leu Gly Arg Ala Val  
115 120

*Sub  
at  
cont*

Ser Leu Leu Val Thr Leu Thr Ile Gly Trp  
125 130

Pro Met Tyr Leu Ala Phe Asn Val Ser Gly  
135 140

Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr  
145 150

His Pro Tyr Arg Val Arg Leu Leu Ile Tyr  
155 160

Val Ser Asp Val Ala Leu Phe Ser Val Thr  
165 170

Tyr Ser Leu Tyr Arg Val Ala Thr Leu Lys  
175 180

Gly Leu Val Trp Leu Leu Cys Val Tyr Gly  
185 190

Val Pro Leu Leu Ile Val Asn Gly Phe Leu  
195 200

Val Thr Ile Thr Tyr Leu Arg Val His Tyr  
205 210

Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly  
215 220

Ala Leu Ala Thr Met Asp Arg Asp Tyr Gly  
225 230

Ile Leu Asn Lys Val Phe His His Ile Thr  
235 240

Asp Thr His Val Ala His His Leu Phe Ser  
245 250

Thr Met Pro His Tyr His Leu Arg Val Lys  
255 260

Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp  
265 270

Asp Thr Pro Phe Tyr Lys Ala Leu Trp Arg  
275 280

Glu Ala Arg Glu Cys Leu Tyr Val Glu Pro  
285 290

*Sub  
al  
Cent*

Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr  
295 300

Trp Tyr Arg Asn Lys Tyr Leu Arg Val  
305

(2) INFORMATION FOR SEQ ID NO:9

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 302 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile  
5 10

Ala Phe Cys Leu Tyr Tyr Val Ala Thr His  
15 20

Tyr Phe His Leu Leu Pro Gly Pro Leu Ser  
25 30

Phe Arg Gly Met Ala Ile Tyr Trp Ala Val  
35 40

Gln Gly Cys Ile Leu Thr Gly Val Trp Val  
45 50

Val Ala Phe Ser Asp Tyr Gln Leu Leu Asp  
55 60

Asp Ile Val Gly Leu Ile Leu His Ser Ala  
65 70

Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr  
75 80

Ser His Arg Arg His His Ser Asn Thr Gly  
85 90

Ser Leu Glu Arg Asp Glu Val Phe Val Pro  
95 100

Lys Val Ser Lys Tyr Leu Asn Asn Pro Pro  
105 110

Gly Arg Val Leu Thr Leu Ala Val Thr Leu  
115 120

*sub a' mt*

Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu  
125 130

Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe  
135 140

Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile  
145 150

Tyr Ser Val Ile Ser Asp Ala Gly Val Leu  
155 160

Ala Val Val Tyr Gly Leu Phe Arg Leu Ala  
165 170

Met Ala Lys Gly Leu Ala Trp Val Val Cys  
175 180

Val Tyr Gly Val Pro Leu Leu Val Val Asn  
185 190

Gly Phe Leu Val Leu Ile Thr Phe Leu Gln  
195 200

His Thr His Val Ser Glu Trp Asp Trp Leu  
205 210

Arg Gly Ala Leu Ala Thr Val Asp Arg Asp  
215 220

Tyr Gly Ile Leu Asn Lys Val Phe His Asn  
225 230

Ile Thr Asp Thr His Val Ala His His Leu  
235 240

Phe Ser Thr Met Pro His Tyr His Ala Met  
245 250

Glu Ala Thr Val Glu Tyr Tyr Arg Phe Asp  
255 260

Glu Thr Pro Phe Val Lys Ala Met Trp Arg  
265 270

Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro  
275 280

Asp Gln Ser Thr Glu Ser Lys Gly Val Phe  
285 290

*Sub  
A  
Cont*

Trp Tyr Asn Asn Lys Leu Ala Met Glu Ala  
295 300

Thr Val

(2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Gly Ala Gly Gly Arg Met Thr Glu Lys  
5 10

Glu Arg Glu Lys Gln Glu Gln Leu Ala Arg  
15 20

Ala Thr Gly Gly Ala Ala Met Gln Arg Ser  
25 30

Pro Val Glu Lys Pro Pro Phe Thr Leu Gly  
35 40

Gln Ile Lys Lys Ala Ile Pro Pro His Cys  
45 50

Phe Glu Arg Ser Val Leu Lys Ser Phe Ser  
55 60

Tyr Val Val His Asp Leu Val Ile Ala Ala  
65 70

Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile  
75 80

Pro Ala Leu Pro Ser Pro Leu Arg Tyr Ala  
85 90

Ala Trp Pro Leu Tyr Trp Ile Ala Gln Gly  
95 100

Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp  
105 110

Val Val Gly Leu Val Leu His Ser Ser Leu  
115 120

*Sub A1  
center*

Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser  
125 130

His Arg Arg His His Ser Asn Thr Gly Ser  
135 140

Leu Glu Arg Asp Glu Val Phe Val Pro Lys  
145 150

Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro  
155 160

Tyr Val Tyr Asn Asn Pro Val Gly Arg Val  
165 170

Val His Ile Val Val Gln Leu Thr Leu Gly  
175 180

Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser  
185 190

Gly Arg Pro Tyr Pro Arg Phe Ala Cys His  
195 200

Phe Asp Pro Tyr Gly Pro Ile Tyr Asn Asp  
205 210

Arg Glu Arg Ala Gln Ile Phe Val Ser Asp  
215 220

Ala Gly Val Val Ala Val Ala Phe Gly Leu  
225 230

Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp  
235 240

Trp Val Val Arg Val Tyr Ala Val Pro Leu  
245 250

Leu Ile Val Asn Ala Trp Leu Val Leu Ile  
255 260

Thr Tyr Leu Gln His Thr His Pro Ser Leu  
265 270

Pro His Tyr Asp Ser Ser Glu Trp Asp Trp  
275 280

Leu Arg Gly Ala Leu Ala Thr Met Asp Arg  
285 290

*Sub  
Alt  
Cont*

Asp Tyr Gly Ile Leu Asn Arg Val Phe His  
 295 300  
 Asn Ile Thr Asp Thr His Val Ala His His  
 305 310  
 Leu Phe Ser Thr Met Pro His Tyr His Ala  
 315 320  
 Met Glu Ala Thr Lys Ala Ile Arg Pro Ile  
 325 330  
 Leu Gly Asp Tyr Tyr His Phe Asp Pro Thr  
 335 340  
 Pro Val Ala Lys Ala Thr Trp Arg Glu Ala  
 345 350  
 Gly Glu Cys Ile Tyr Val Glu Pro Glu Asp  
 355 360  
 Arg Lys Gly Val Phe Trp Tyr Asn Lys Lys  
 365 370  
 Phe Xaa

## (2) INFORMATION FOR SEQ ID NO:11

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 224 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Trp Val Met Ala His Asp Cys Gly His His  
 5 10  
 Ala Phe Ser Asp Tyr Gln Leu Leu Asp Asp  
 15 20  
 Val Val Gly Leu Ile Leu His Ser Cys Leu  
 25 30  
 Leu Val Pro Tyr Phe Ser Trp Lys His Ser  
 35 40  
 His Arg Arg His His Ser Asn Thr Gly Ser  
 45 50

*Sub A1  
cont*

Leu Glu Arg Asp Glu Val Phe Val Pro Lys  
55 60

Lys Lys Ser Ser Ile Arg Trp Tyr Ser Lys  
65 70

Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met  
75 80

Thr Ile Ala Val Thr Leu Ser Leu Gly Trp  
85 90

Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly  
95 100

Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr  
105 110

Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg  
115 120

Glu Arg Ile Glu Ile Phe Ile Ser Asp Ala  
125 130

Gly Val Leu Ala Val Thr Phe Gly Leu Tyr  
135 140

Gln Leu Ala Ile Ala Lys Gly Leu Ala Trp  
145 150

Val Val Cys Val Tyr Gly Val Pro Leu Leu  
155 160

Val Val Asn Ser Phe Leu Val Leu Ile Thr  
165 170

Phe Leu Gln His Thr His Pro Ala Leu Pro  
175 180

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu  
185 190

Arg Gly Ala Leu Ala Thr Val Asp Arg Asp  
195 200

Tyr Gly Ile Leu Asn Lys Val Phe His Asn  
205 210

Ile Thr Asp Thr Gln Val Ala His His Leu  
215 220

*Part A  
Cont*

Phe Thr Met Pro

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCTCTTTGT GCGCTCATTC

20

(2) INFORMATION FOR SEQ ID NO:13

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGTACCAAGA AAACGCCATTG

20

(2) INFORMATION FOR SEQ ID NO:14

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TAYWSNCAYM GNMGNCAYCA

20

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

RTGRTGNGCN ACRTGNGTRT C

21

*Sub  
at  
Cont*